Complex visual data analysis, uncertainty, and representation

Christian D. Schunn

Lelyn D. Saner

Susan K. Kirschenbaum

J. Greg Trafton

Eliza B. Littleton

In press. To appear in M. C. Lovett & P. Shah (Eds.), Thinking with Data. Mahwah, NJ: Erlbaum.

Report Documentation Page			*Form Approved* *OMB No. 0704-0188*
Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.			
1. REPORT DATE **2007**	2. REPORT TYPE		3. DATES COVERED **00-00-2007 to 00-00-2007**
4. TITLE AND SUBTITLE **Complex visual data analysis, uncertainty, and representation**			5a. CONTRACT NUMBER
			5b. GRANT NUMBER
			5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)			5d. PROJECT NUMBER
			5e. TASK NUMBER
			5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Naval Research Laboratory,Navy Center for Applied Research in Artificial Intelligence (NCARAI),4555 Overlook Avenue SW,Washington,DC,20375**			8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)			10. SPONSOR/MONITOR'S ACRONYM(S)
			11. SPONSOR/MONITOR'S REPORT NUMBER(S)
12. DISTRIBUTION/AVAILABILITY STATEMENT **Approved for public release; distribution unlimited**			
13. SUPPLEMENTARY NOTES			
14. ABSTRACT **How do problem solvers represent visual-spatial information in complex problem solving tasks? This paper explores the predictions of symbolic computation, embodied problem solving and a neurocomputational theory for what factors influence internal representation choices. Across two studies, data are collected from experts and novices in three different, complex visual-spatial problem-solving domains (weather forecasting, submarine target motion analysis, and fMRI data analysis). Internal spatial representations are coded from spontaneous gestures made during cued-recall summaries of problem solving activities. Analyses of domain differences, expertise differences, and changes over time with problem solving suggest that neurocomputational constraints play a larger role than the nature of the visual input or the nature of the underlying real world being examined through problem solving, especially for expert problem solvers. The particular neurocomptuational feature that was found to drive internal representation choice is the required spatial precision of the main goals of problem solving.**			
15. SUBJECT TERMS			

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**53**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

Abstract

How do problem solvers represent visual-spatial information in complex problem solving tasks? This paper explores the predictions of symbolic computation, embodied problem solving and a neurocomputational theory for what factors influence internal representation choices. Across two studies, data are collected from experts and novices in three different, complex visual-spatial problem-solving domains (weather forecasting, submarine target motion analysis, and fMRI data analysis). Internal spatial representations are coded from spontaneous gestures made during cued-recall summaries of problem solving activities. Analyses of domain differences, expertise differences, and changes over time with problem solving suggest that neurocomputational constraints play a larger role than the nature of the visual input or the nature of the underlying real world being examined through problem solving, especially for expert problem solvers. The particular neurocomptuational feature that was found to drive internal representation choice is the required spatial precision of the main goals of problem solving.

Representations and Complex Problem Solving

A core thesis of cognitive science is that representations, be they structures inside the head of the problem solver (internal representations) or structures in the environment of the problem solver (external representations), are fundamental to understanding problem solving behavior (Markman, 1999). There may be some debate about the underlying nature of these representations (e.g., the relative weight that problem solvers place on internal versus external representations, whether the representations are symbolic or not), but all cognitive scientists endorse some form of underlying representation driving behavior (Dietrich & Markman, 2000).

The value of talking about representation is that computation (the definitional core of cognition in cognitive science) can only be meaningfully defined over some type of representation. Computation at its root consists of a data structure (for input, output, and perhaps something being stored in between) and some process. One cannot talk about the process without describing the data structure. More importantly, different data structures enable certain computations to be done easily, whereas other data structures support other computations. Thus, the choice of data structure (representation) helps explain why a problem solver does or does not successfully engage in a given process (cognition/behavior) or perhaps why a process takes as long or as short as it does.

The goal of this chapter is to argue that representation can and should be studied directly rather than be ignored or left as an explanatory variable. We begin with a discussion of how representations can be measured. We then examine several of the main theoretical paradigms of cognitive science to see what *predictions* they make for representation choice. Finally, we present data from two studies designed to test these predictions.

Measuring Representations

While the theoretical construct of representations are fundamental to cognitive science theorizing, the dirty secret of cognitive science is that we do not have a good theory for predicting what representations a given problem solver in a given situation will use. The most common treatment of representation is as an explanatory, intermediate (i.e., hidden) variable of performance. In other words, we observe inputs and outputs, and perhaps infer process via a description of intermediate states, but the representation is something only posited or assumed in order to explain the relationship between inputs and outputs.

To help explain why we do not have theories of how problem solvers pick a representation, we unpack in this section the basic measurement problem for internal and external representations. On some occasions, there is some attempt to measure the representations used by the problem solvers, to verify the assumptions that were made about the representations. Measuring representations is no easy feat. Even with external representations, the measurement task is challenging. Different observers can 'see' different things in the world around them. Humans are fundamentally limited in how much information they can attend from their perceptual input (Treisman, 1969). Thus, we do not know what information in the environment to which they are attending. In a complex problem-solving situation, the external environment can be quite complex, and thus there can be a very large amount of information that is not attended. Then there are the interactions of human perception with displays. Some perceptual features are salient and easily encoded (e.g., strong color contrasts, motion onset, object appearance) whereas other perceptual features not salient or not so easily encoded (e.g., conjunctions of features, absolute pitch of sounds, etc.). Operations are easier on some external

representations than on other external representations, and cognitive science talks about this interaction as affordances (Neisser, 1976).

So, to capture what external representations a problem solver is using, we need to know what objects in the environment and what features of the objects are being attended, rather than just looking at what perceptual input is available in the environment and how it is structured. To some extent, we can measure what visual objects are being encoded through eye-tracking studies—people encode primarily what objects they foveate, and they tend to foveate objects they are thinking about (Just & Carpenter, 1976a, 1976b). But, people do encode some information from parafovea, and they sometimes do not encode information while staring blankly at the environment while their mind is elsewhere.

Another approach to measuring external representations is a mixture of ethnographic or quantitative observation and interviews (Hutchins, 1995a). The external objects that problem solvers name or discuss while problem solving provide some good clues to what objects and features of the external environment are being attended. However, linguistics manifestations of the perceptual inputs in conversation may be systematically changed or selective, and may exclude some of the perceptual inputs that were attended. Thus, measuring external representations continues to be a challenging task.

Even more challenging is the task of measuring internal representations. Even when information is represented using a fairly perceptual-based internal representation, people are capable of recoding information from one perceptual code into another. For example, people given words visually will often transform the words into an auditory code (thereby producing auditory confusion errors rather than visual confusion errors). More problematic are recodings of input into more symbolic, abstract forms. People can build mental links between objects that aid

later retrieval (Altmann & Trafton, 2002), develop new groupings of objects (called chunks) allowing them to simultaneously represent more information in mind, and categorize objects in ways unrelated to their perceptual input (e.g., the categories of 'even numbers', 'uncles', and 'abstract categories').

Internal representations can only be measured indirectly, if 'measured' is even the correct term. The basic problem is that we must typically look for some kind of external manifestation of the internal representation and that external manifestation may be only a distant echo or heavily distorted copy of the internal representation. For example, we can ask a problem solver to draw a picture or verbally describe a given situation. The drawing may omit verbal representations, the verbal description may miss visual/spatial descriptions, both may be incomplete descriptions of those sensory-based representations, and both may be missing more abstract representations. Moreover, the act of having to describe the situation may push the problem solver to represent features they would not have otherwise mentioned, but only looking at spontaneous produced external manifestations (e.g., spontaneous speech, gesture, or drawings) may only capture communicatively-focused representations.

One approach to measuring internal representations more directly is to use neuropsychological techniques such as neuroimaging (MEG, fMRI, ERP, or PET) or single-cell recording. While this approach holds some promise for more direct measurement, the methodological challenges of these methods (e.g., spatial and temporal resolution, extremely high noise levels) prevent their use in most complex problem solving situations, and currently we know relatively little about the internal codes of the brain (i.e., what activation in different brain areas even means).

The reason that we discussed the basic measurement problem for internal and external representations is to help explain why we do not have theories of how problem solvers pick a representation. The answer is that we rarely go to the trouble of measuring representations, and when we do, we are not so sure we have done it correctly. Thus, the set of results or phenomena related to representation choice is rather thin, and this would explain why we do not have much theorizing about representation choice.

Theorizing about Representation Choice

While there are relatively few theories of representation choice (i.e., how problem solvers choose representations) in cognitive science relative to how important representation choice is to cognitive science, there are some theories (e.g., Kaplan & Simon, 1990; e.g., Lovett & Schunn, 1999). This theories can be summarized under the heading of Search and Rational Choice: problem solvers consider different representations when they are unsuccessful and select the representations that turn out to lead to more successful problem solving. There is also an assumption that problem solvers start with certain salient features, but there is no theoretical specification of what features will be salient.

There is, however, considerable theorizing about the consequences of different representation choices. By applying this same general idea of rational choice, this broad theorizing about consequences can be turned into theories of choice. Specifically, with some form of rationality or asymptotic optimality assumptions, one could predict that problem solvers will tend to select the more useful or optimal representations, especially with expertise. While rationality and optimality are still somewhat controversial theoretical assumptions, they are not so controversial when 1) we limit ourselves to directional optimality (i.e., better is more likely than worse), and 2) we focus on the case of expert behavior. Of course, experts also likely to be limited to some

form of bounded rationality (Simon, 1956), and it is possible for experts to get "stuck" using a representation chosen early but that becomes relatively successful through high levels of practice.

In terms of consequences of representation choice, we believe there are three broad theoretical camps in cognitive science: symbolic problem solving, embodied problem solving, and neurocomputational problem solving. If we apply this rationality assumption to these different theoretical camps, we can derive different predictions about representation choice. The goal of this chapter is to spell out these predictions, and explore their relative usefulness for predicting representation choice in several complex problem-solving domains in which experts work with highly uncertain visual data.

Representations in Symbolic Problem Solving

Cognitive science began with a focus on the internal and a focus on symbol processing. This approach was heavily centered in Pittsburgh and was heavily influenced by Herb Simon (Chi, Feltovich, & Glaser, 1981; Kaplan & Simon, 1990; Klahr & Dunbar, 1988; Kotovsky, Hayes, & Simon, 1985; Larkin, McDermott, Simon, & Simon, 1980; Lovett & Schunn, 1999). The assumption of this symbolic problem solving camp is that representation is determined by task structure (in their language, the problem space) modulo various memory limitations. More specifically, experts are assumed to have representations that strip away irrelevant details and focus upon or build features that highlight the deep structure of the task at hand. For example, in a classic study, Chi, Feltovich, and Glaser (1981) found that physics experts represented simple physics problems not in their superficial terms (incline planes, pulley problems, etc), but in terms of their solution types (conservation of energy problems, F=ma problems, etc).

In terms of the constraints from the external world, the symbolic problem solving approach does not say much. Internal representations must be derivable in some ways from the external input, and it is assumed that people incur some processing costs for encoding and transformation. But there is not a detailed or higher-level theory for how the external world would shape the representational choice of an expert.

In terms of constraints of the internal world (i.e., the brain), the symbolic problem solving approach places even fewer constraints as a general theoretical approach. When one gets into particular theories, memory constraints may become apparent. For example, expertise, in some accounts, is not thought to change the number of chunks a person can hold onto in memory. However, the size of chunks is not theoretically constrained, and some theories posit that experts can learn higher order structures which allow them to circumvent the number of chunks constraint as well (Gobet & Simon, 1996). Thus, even the memory constraint is not very constraining in terms of predicting expert representation choice.

In sum, although the symbolic problem solving approach assumes that representation choice is one of the most important issues in predicting behavior, it says relatively little about how individuals make that choice. At most, the symbolic camp makes the general prediction that people start with representations that focus on salient or superficial features and move to representations that somehow better facilitate task-relevant computations (Lovett & Schunn, 1999).

Representations in Embodied Problem Solving

Beginning in the late 1970s, some researchers began to look to the structure of the external environment, and more specifically, the ways in which the human body fit into the structure of the external environment to explain human cognition and performance (Gibson, 1979; Neisser,

1976), rather than focusing so heavily on logical task structure and internal symbol manipulation. Several different lines of theorizing developed under this approach, including situated cognition (Suchman, 1987), distributed cognition (Hutchins, 1995a, 1995b), embodied action (Fu & Gray, 2004; Gray, John, & Atwood, 1993), and perceptual symbol systems (Barsalou, 1999). For example, Ed Hutchins showed in great detail how the performance of an airplane pilot depends upon being able to apply simple perceptual heuristics to correctly laid-out instruments.

An important construct in many (but not all) lines in this general approach is the notion of affordances. A display, instrument, or artifact is thought to afford certain actions and cognitions and not afford others, meaning the combination of past experiences and body/mind structure make certain actions very easy to do with a given display, instrument, or artifact. Along these lines, one might argue that experts adopt internal representations that maximize the affordances relationship between common tasks and external objects used to execute those tasks. Yet, without a very formal description of what affordances are and a predictive framework for specifying what affordances come from what tasks, the affordances framework tends to be circular in nature. Affordances are said to exist in an object when we observe smooth performance with that object. In other words, affordances are used to both describe and explain referentially to one another.

A different, more straightforward prediction of representation choice from the embodied problem solving approach is to predict an internalization of the external. If cognition and problem solving is so heavily driven by the external world, then internal representations must be closely tied to the external world. That is, the choice of internal representation must be heavily driven by the nature of the external input (e.g., if input is 2-dimensional, then internal representations are likely to be 2-dimensional, though cf. Scaife and Rogers (1996)).

A related prediction focuses on the action rather than input side of performance. Sometimes input is in one form, but our actions occur in a different form. For example, the input a submariner gets from Sonar is very 2-dimensional, but the actions they take move the submarine in 3-dimensions. The expert problem solver may adopt an internal representation that is close to the external input form, thereby reducing the complexity of computations going from input to cognition. Alternatively, the expert problem solver may adopt an internal representation that is close to the external action form, thereby reducing the complexity of computations going from cognition to action. Thus, one could also predict from the embodied problem solving perspective that choice of internal representations are driven by the external reality of the action space that the problem solver is working within.

In sum, the embodied problem solving perspective predicts a close match in internal representation choice to either the input representations or the output environment.

Representations in Neurocomputational Problem Solving

While the abstract and embodied problem solving perspectives are themselves very different, they share a belief about the very adaptable nature of human representations—at some level, any internal representation is possible, given an appropriate task and/or appropriate external input. Or said another way, at some level, the details of the human brain are irrelevant to predicting the range of possible expert internal representations or the choice among possible representations.

At the same time, we now know a great deal about the way the human brain processes perceptual information. Since the early 1960s, we have known that the brain has very specific maps of perceptual information (Hubel & Wiesel, 1962). Since the 1980s, we have known that the brain has elaborate pathways by which complex representations of perceptual information are gradually and carefully constructed (Ungerleider & Mishkin, 1982). Whiles these pathways must

be rebuilt within each person through development, the general structure is very similar from individual to individual. The structure, therefore, is primarily driven by a genetic code, or at least a genetic code interacting with general features of the overall environment that are the same for everyone. Therefore, the structure of these perceptual pathways is not likely to vary significantly across individuals with different areas of expertise.

Is the way that humans process perceptual information relevant to internal representation structure? In the 1970s and 1980s, there was considerable debate about whether mental imagery was entirely symbolic or more analog in format, and about whether it relied entirely on higher level cognition or whether the basic perceptual machinery was involved (e.g., Kosslyn, Ball, & Reiser, 1978; Kosslyn, Pinker, Smith, & Shwartz, 1979; e.g., Pylyshyn, 1973; Pylyshyn, 1981). By the 1990s, most researchers considered the issue resolved with the introduction of neuroscience evidence that showed clearly that mental imagery depended heavily upon all the same neural structures as did perception, except the very earliest part of perceptual pathways (Kosslyn, 1994; Kosslyn et al., 1999; Kosslyn, Thompson, Kim, & Alpert, 1995), although there is still some ongoing debate (Pylyshyn, 2002).

In sum, human brains have complex information processing pathways that process information in very particular ways and are common across individuals regardless of expertise, and those processing pathways are used for at least some internal representations. From that point of view, it seems plausible that the details of our neurobiology would influence the range of possible internal representations that an expert could have. This restriction of choices is particularly relevant for representations that have a perceptual or analog character rather than purely symbolic flavor, in part because we understand in great detail the perceptual systems in the brain but understand very little about the symbol manipulation systems in the brain. On the

other hand, the cognitive architectures approach to modeling cognition (e.g., Anderson & Lebiere, 1998; Kieras & Meyer, 1997; Newell, 1990) also support the general idea that the underlying cognitive architecture places some restrictions, albeit relatively weak thus far (Anderson, 1990), on what kinds of representation systems are possible. Thus, there may also be neurocomputational restrictions of choices of more symbolic representations as well.

Because we have now wandered into the realm of biology, it may be worth bringing forward a framework from biology for understanding how new representations come about. In biology, the new representations are external physiological changes in a species, or creations of new species. In our case, the new representations are new internal representations with developing expertise. The symbolic problem-solving framework corresponds to the biological notion of adaptation: new representations are developed by adapting existing representations to the current task structure (Kaplan & Simon, 1990; Schunn & Klahr, 1996, 2000). The embodied problem-solving framework corresponds to notions of an analog (rather than homolog). Analogs are structures that arise from a different evolutionary source but serve similar functions, whereas homologs are structures that arise from similar evolutionary sources. The internal representation in the embodied problem-solving framework is thought to be at some level a copy of input or output representations, selected from a different neural substrate to serve a similar function (i.e., an analog). By contrast, the neurocomputational problem-solving framework corresponds to notions of exaptation (Gould & Vrba, 1982). Under this account, internal representations that were developed over evolutionary time for one set of tasks can become co-opted or exapted to a new use as new tasks occur. To be more specific, the human problem solver is born with internal representational abilities that were there to support very traditional tasks shared with other mammals (e.g., object recognition, object manipulation, navigation, etc). The human problem

solver must make use of those fixed set of representational abilities to build representations for the range of modern tasks that humans now become expert in.

Following this line of argument further, we can then move to understanding the influence of neurocomputational constraints on the choice of particular representations for a particular task, not just on the set of possible representations. The trick is to focus on notions of efficiency or affordances, as do the abstract and embodied problem solving frameworks. Different neuropsychological representational systems represent information in different ways in order to support different tasks (Previc, 1998; Ungerleider & Mishkin, 1982), implying that some computations are accurately or more quickly performed with some representational systems than with others. Therefore, as with exaptation in biology, we can predict that expert problem-solvers will tend to select the internal representation system whose neurocomputational abilities best support the expert's task at hand. For example, if the task requires the expert to represent themselves at the center of a full 360 degree space of mental objects placed around them, and if only one neural system supports representations in the full 360 (rather than just a frontal 120 or 180), then this approach would predict in a rather straightforward fashion that the problem solver would use that neural system for internal representations of this task. We will say more about different possible human neural systems and their neurocomputational abilities in a later section.

Comparison of Representational Predictions

Table 1 presents a comparison of the general predictions made about internal representation under the three theoretical camps. All three camps agree that affordances should matter in that experts will choose internal representations that best match the cognitions that need to be performed, and that different representations have different affordances. At some level, all three camps agree with the basic characterization provided originally by the symbolic camp that the

story of affordances is best cast in computational terms—affordances reduce necessary computations by the problem solver.

The camps do differ in exactly how the affordances are described. More specifically, they differ in the objects against which affordances are primarily defined. This focus brings us to the second dimension of comparison, the issue of whether the external world matters. The symbolic camp is somewhat neutral on this point. The external world may or may not influence internal representation choice, depending upon whether there are features of the external world that are particularly helpful. In other words, if the structure of the external world is not useful for the problem solver, then the problem solver may choose to work entirely within an internally constructed representation that has little to no relationship to the external world. One can point to characterizations of insight problems in these terms: one core trick in solving the insight problem is to move away from the salient details of the external world and develop a new representation (Kaplan & Simon, 1990; Perkins, 1994).

By contrast, the embodied problem solver camp predicts that the external world will have a strong role in influencing internal representations. The reason is that the embodied problem solving perspective assumes that experts organize their external worlds such that they can make heavy use of the external world to guide their problem solving (Hutchins, 1995a). In other words, there is a belief that real world tasks are typically embedded in complex socio-technical systems that are influenced by the individual expert problem solver (in which parts of their rich environment they chose to use) and by collections of expert problem solvers (in influencing the construction of artifacts). Expert problem solvers thereby make it possible for their internal representations to have a close affinity to the external world around them, simplifying the translation between internal and external, and yet still have very successful problem solving.

The neurocomputational problem solver chooses a more nuanced and complex stance on the role of the external world on internal representation choice. The human perceptual system involves a division and modulation of separate perceptual features along with some integration across perceptual modalities. For example, vision can be processed separately from sound, and even within vision, color can be processed separately from orientation, and object identity can be processed separately from object location. At the same time, the brain can also integrate across very different senses, building, for example, a spatial map of the environment from visual, auditory, and tactile cues (Previc, 1998). Attention adds another layer, by being able to reduce or even remove the processing of certain streams of information (Broadbent, 1957; Pylyshyn, 1994; Treisman, 1969). The bottom line is that the neurocomputational perspective assumes that the external world has a strong influence on internal representation choice because our internal representational machinery makes heavy use of perceptual processing systems, but that the problem solver has the ability to ignore certain perceptual inputs. Thus, only perceptually segmentable aspects of the external environment that need to be processed for the task at hand will influence internal representation choice. The perceptually segmentable constraint on what can be treated separately depends upon the neurophysiological limits of our perceptual system. What is segmentable is a complex story that we cannot fully unpack here, but it is sufficient for our purposes here to say that some features can be processed separately whereas others cannot (Wolfe, 1994).

The final dimension of comparison is the space of possible choices of internal representations. For the symbolic and embodied problem solving camps, essentially anything, in theory, is possible. For the symbolic problem solving perspective, the set of likely choices are likely to be mostly symbolic in one way or another, although a mixture of symbolic and analog is

possible (Larkin & Simon, 1987; Tabachneck-Schijf, Leonardo, & Simon, 1997). For the embodied problem solver perspective, the choices are obviously heavily influenced by the external world, but essentially anything in the external world could be mimicked internally, at least in theory. The perspective that is most distinctive on this dimension is the neurocomputational problem solving perspective. The neurocomputational perspective holds that the problem solver can only use a very fixed set of representational schemes. This fixed set is instantiated as human brain systems and is heavily determined by evolutionarily important tasks.

Testing the Theoretical Camps

No simple set of experiments can easily test between very different theoretical paradigms because of the all the additional assumptions required to account for a particular experimental situation. However, we can ask how useful the different paradigms are for explaining internal representational choice in a few cases. In this chapter, we describe two studies designed to look at internal representations of experts, and the situations of these experiments were chosen such that the different theoretical camps would make different concrete predictions for internal representation choice. In particular, we examined representation choice in how experts deal with uncertainty while analyzing complex visual/spatial data. We realize that we cannot generalize from these studies to the utility of the different theoretical camps overall. However, these studies do provide a concrete example of how one can empirically test between the utility of the different paradigms.

Both studies examine one very particular aspect of representation: how people represent visual/spatial information. The world is 3-dimensional, but most information sources that experts in complex domains interact with are 2-dimensional (e.g., paper and computer screens). The world exists relative to the problem-solver in egocentric terms, but information sources often

present visual/spatial data in exocentric terms. The world is life-sized (again by definition), but expert information sources often present scaled versions, either much larger (e.g., via microscopes) or much smaller (e.g., satellite images). Given this diversity of reality and input, how will the problem solver represent their problem solving states internally?

The symbolic camp tells us to conduct a task analysis. Find out what strategies and representations are possible, and which are most efficient for the task at hand. The embodied problem solving camp suggests that representations will match either the form of the external input or the external reality of the problem. What about the neurocomputational problem solver? Here the devil is in the details—in order to develop predictions, we need to select an account (among several competing accounts) for how the brain represents visual/spatial information. We have selected the ACT-R/S theory, and explain it with just enough detail so that the predictions can be made for our current needs.

Brief Overview of ACT-R/S

ACT-R/S (Harrison & Schunn, 2001) is a neurocomputational theory of the visual/spatial representational and computational abilities of the human mind. It integrates current neuroscientific understanding of how the human brain represents visual/spatial information into the ACT-R 5.0 (Anderson, Bothell, Byrne, & LeBiere, 2002) view of how the mind achieves complex problem solving through a rich mixture environment encoding, memory retrievals, and skill applications through goal-directed behavior. In particular, ACT-R/S posits that there are three different visual/spatial representations (see Figure 1), which we call buffers. The three representations make use of different neural pathways, tend to get used for different kinds of basic perceptual/motor tasks, have fundamentally different ways of representing space, and have

different strengths and weaknesses. Note that these buffers are multimodal in that they integrate spatial information coming from vision, audition, touch, locomotion, and joint sensors.

The first representation is the Visual Buffer. It is used for object identification and represents information primarily around the region that they eyes are attending to, and represents information in approximate shape terms and approximate size and location. Historically, this buffer has been called the "What" visual pathway. Its representation of the world is primarily a 2-dimensional world, with objects occupying space in the fronto-parallel plane (i.e., like on a computer screen or chart on the wall in front of you). That is, there are approximate above/below and left/right relationships, but no strong distance and exact orientation information.

The second representation is the Manipulative Buffer. Historically, it has been called the "Where" visual pathway. It is used for grasping objects and tracking of moving of objects, representing information close to within reach, but also all the way around the person. It represents spatial information in highly accurate metric terms, which is required for object manipulation, and in a true 3-D fashion. It is not good at figuring out what objects are, but it knows exactly where they are and what there component shapes are.

The third representation is the Configural Buffer. It is used for navigation in small and large spaces, figuring out where you are, where you want to go, and how to get there. It represents information in terms of egocentric range vectors to blobs (e.g., the desk is approximately so far away, with the left and right side being at such and such angles from me). Locations are configurations of such vectors (e.g., I am at the location that is so far away from the door and such distance from the window, with a given angle between the two).

Complex-Problem Solving, Representation choice, and ACT-R/S

The strong assumption in ACT-R/S is that these three representations are the only representations (other than verbal) that a novice or expert can use for problem solving. In other words, an expert cannot invent a new visual/spatial representation that does not use one (or more) of these three representations, and that there representations will be limited computationally in the same ways as novices based on the properties of these three visual/spatial representation systems. That is, people are assumed to be fundamentally limited by their neurobiology.

ACT-R/S assumes that people can translate between the three representations. In fact, for many tasks, translation and simultaneous activation of different representations is necessary. For example, in order to figure out one's location (a Configural task), one needs to identify what the landmarks are (a Visual task). This ability to translate between representations in general is what makes much of cognitive psychology so difficult because the internal representation can differ dramatically from the input form and can vary substantially across individuals, and the choice of internal representation fundamentally influences performance. For example, people can have visual representations of auditory stimuli, producing visual confusions rather than auditory confusions. In the case of ACT-R/S, a person can take arrangements of distant objects presumably only representable in the Configural space and translate it into a miniature 3D model in the manipulative space, or a flat visual map representation in the Visual space. The way that the person is internally representing the objects will then strongly determine how spatial features are encoded, and thus an important determiner of performance.

The choice of which representation is used will be influenced by input: things in flat displays will tend to start out as Visual; things within reach will tend to start out as Manipulative,

and things out in the distant will tend to start out as Configural. However, the choice of representation will also be influenced by functional factors. ACT-R, the parent theory, assumes that people make procedural choices on the basis of past experiences of success and amount of effort with the choices. In other words, it predicts that people will tend to select choices that led more often in the past to successful attainment of goals, but also taking into account how much effort (primarily in amount of time) was required to achieve those goals. There are more formal, mathematical instantiations of the choice process and the learning of preferences, but the general understanding of this point will suffice for here. ACT-R/S, then, assumes that people will tend to move towards representations that have been generally more functional for the goal task at hand. Because the three different representations have very different basic representational form and computational abilities, the match of representation to task should be a strong influence on representation choice. Because this choice preference is embedded in a learning theory, the prediction is that this preference for a particular representation will be more pronounced with increasing expertise in a task.

Uncertainty Predictions from ACT-R/S

With all that theoretical background on ACT-R/S and how it might apply to complex problem solving, we can now come full circle back to the issue of visual/spatial representations in complex problem solving with uncertainty. The three different spatial systems have varying degrees of match to spatial certainty. *All things being equal*, ACT-R/S then predicts that problem solving, especially in disciplines with complex visual displays, will vary as a function of spatial certainty levels of the scientist doing the data analysis: Manipulative representations will be used when spatial certainty levels are the highest because the Manipulative space represents spatial location and features in very precise terms; Visual representations will be used when spatial

certainty levels are the lowest because the Visual space represents spatial location and features in very approximate terms; and the Configural representation sits somewhere in between, with precise angles, but approximate distance and very approximate shape information.

Of course, all things are not often precisely equal. Input of information will come in a particular form. The particular goals of the data analysis will influence the functional relevance of different representations, as well. Expertise will play a role here, too, as experts may be more sensitive to functional relevance and less sensitive to initial input form.

In sum, ACT-R/S makes a variety of predictions for how experts will represent visual/spatial information during data analysis, and one of those predictions involves relative uncertainty levels. We thought of this uncertainty prediction as a very novel prediction to the psychology of problem solving, in clear contrast to the predictions of the symbolic and embodied problem solving camps. The symbolic problem solving framework makes relatively few predictions about internal representation choice, and the embodied problem solving framework predicts a match of internal representations to either input or action external representations; neither make a predictions about the relationship of internal representation choice and uncertainty levels. We examine two studies of complex problem solving in a several domains to see which perspective could successfully predict (not just explain) observed (although somewhat indirectly by necessity) internal representation choices.

Study 1: Expert/Novice Comparisons in a Traditional Submarine Task

Overview

This study examined expert, intermediate, and novice representations of 3-Dimensional space while solving the complex spatial task of finding an enemy submarine using a simplified computerized environment of a traditional submarine sonar setup. We carefully examine

participant spontaneous gestures as an indicator of how they are internally representing spatial locations during problem solving.

Participants

In this study, 16 submarine officers total participated: six students, six instructors, and four commanders. The students were recent graduates of the Submarine School's Submarine Officers' Basic Course (SOBC). The instructors were Junior Officers who were teaching those courses at the time of the study. The commanders were Commanding Officers (COs) and Executive Officers (XOs), three of whom were active-duty and one who was retired. In the US Navy, the most expert individuals are considered too valuable to spend time teaching, and thus the instructors are the intermediate level participants.

Procedure

The procedure involved two simulated scenarios in Ned, a simulation environment built in a previous project for studying the expertise of determining a solution (see Materials). First, the participant was familiarized with the ways to gather information about potential contacts in the simulation environment. Then the participant was asked to think aloud as he solved the problem. Each officer worked for approximately 20 minutes to determine the location of an enemy submarine (called a solution). Once a solution was found, the experimenter initiated a retrospective interview. This procedure of problem solving and retrospective interview was then repeated for a second scenario.

During the retrospective interview, the participant gave a general summary of the scenario. Next, the participant was cued to explain specific moments in the simulation he just completed. Cued by predetermined screen shots or short clips of the screen at different moments in the scenario, he was asked to talk about what he was thinking, what the problem was that he was

addressing, and what happened just after this moment. The participant responses were video-taped. The experimenter asked the participant to view the screen once, and then once he was ready to answer, turn away from the screen to speak to the experimenter. This physical manipulation of the screen and the participant was intended to ensure that the participant used hand gestures when he wanted to convey spatial elements and not vague points or gestures to the screen to convey explanations. In addition to the preset screen shots and clips, we generated questions opportunistically during a session, for example, when we wanted to clarify a participant's explanation.

Materials

Ned is a small-scale submarine control room simulation (Ehret, Gray, & Kirschenbaum, 2000). While it provides the functionality to perform all the functions necessary to locate a contact, all of the data on the contact are simulated. They are not represented by a high-fidelity model, but rather by noise plus the true values for key parameters. The interface that Ned uses is a composite of common submarine display content without being a copy of any specific deployed system. As it is generic to all contemporary systems, submariners will be familiar with these displays and their functionality.

Ned was developed with four scenarios, two of which were randomly assigned to each participant. All scenarios have two contacts—a rather noisy merchant and a quieter enemy submarine. In two of the scenarios, the subsurface contact is moving at a constant course and speed and in the other two it is maneuvering every ten minutes, on average. The merchant ship appears at the beginning of the scenario, and after about one minute the submerged contact appears. In some scenarios, when the sub appears, it is dead ahead of own-ship, necessitating a speedy maneuver on the part of own-ship to avoid the possibility of a collision and get into a

more advantageous position relative to the sub. In the other scenarios, the submerged contact appears ahead of own-ship, but in not as dangerous a position, still requiring own-ship to maneuver but not as quickly as in the dead-ahead scenarios. Eventually, the submerged contact drives toward the merchant and trails it, giving the impression that the sub is threatening the merchant. Also, as the scenario progresses, the spatial relationships of the two contacts become complicated and critical as the two ships get closer to one another.

Figure 2 presents two sample screen shots from Ned. The left half of the top screen shot presents a diagram showing the presence of certain sound frequencies[1] at different angles of input. The right half shows information on different sound 'tracks' that the problem solver has chosen to follow. The bottom screenshot shows a geosituational view.

Predictions

Note that none of the input in Ned shows the equivalent of a view out of a window although there is a bird's-eye-view with own ship in the center and lines of bearing to other platforms, and current solution, if available. The visual/spatial displays are all 2-dimensional, complex displays. At the same time, the real world being reasoned about is a very, very large, 3-dimensional world. How will problem solvers represent this situation internally?

The symbolic perspective predicts that problem solvers will select whatever representation minimizes mental workload and maximizes accuracy—in this complex task, we had no idea what that would be, and thus felt that no predictions were being made by the symbolic perspective other than whatever internal representation is most correlated with high performance within groups would be more likely to occur in experts. The embodied problem solving perspective predicts that problem solvers will use either 2D display-based reasoning (the input) or large-scale

[1] Because true frequencies are classified, the values used in Ned are made-up and the convention used was explained to the participants during training.

3D (configural) reasoning (the real world). By contrast, the neurocomputational perspective suggests that problem solvers will move from a display or configural representation to a manipulative (small 3D) representation because 1) configural or display representations are more appropriate for weak initial knowledge of location and distance, and 2) manipulative representations are more appropriate when location and distance are more accurately known. The neurocomputational perspective is the only one that very clearly predicts a change in internal representation choice for this task over time.

Gesture Coding

Visual-spatial representations were coded from the spontaneous gestures. Configural gestures were made with the hand or arm such that the fingers are pointing in a direction without attempting to pick up or place or otherwise manipulate imaginary objects. These were usually one-handed gestures and one-dimensional, but some were two-handed when they have a quality of pointing into the distance. They can represent limited motion, for example in a single direction, but again only if it seems the motion is in far-space and not being manipulated in curves and complex dimensions. See Figure 3 for an example of a two-handed configural gesture in which the hands represent the angle at which the target is at relative to the heading of own-ship.

Manipulative gestures placed objects and activity in a nearby space, such that the participant can actually manipulate or place the imaginary objects. Gestures include two-handed gestures showing two contacts and the relative motion involved or changes in bearing and curves in paths or course. Gestures in which the hand-shape suggests placing or holding as opposed to strictly pointing were also coded as manipulative. Figure 4 presents an example in which a student represents the submerged contact in a stationary position and own-ship moving forward and then

turning left to follow behind the other hand (the sub). This gesture represents relative positions, motion and a complex path for own-ship.

Display-based gestures would have been gestures that involved gestures that place objects and activity on a flat surface in the fronto-parallel plane. However, in this study, those kinds of gestures did not occur, and thus are not mentioned further. There were also uncertainty-based gestures, in which participants shrugged or wiggled their hands indicating uncertainty about the situation, but those gestures do not directly indicate spatial representations and thus are not discussed further in this chapter.

Reliability of the coding was between 84% and 92% agreement depending upon the category and was established with a second rater coding a randomly selected 20% of the data. The analyses reported here focus on the gestures made during the first and last maneuvers of both scenarios to show change in representations during problem solving (in addition to changes with expertise).

It is important to note that spontaneous gestures are an indirect measure of internal representation, and that they are likely to have biases as such a measure. For example, the gestures may be influenced by communication goals (McNeill, 1992). However, this measure of internal representation is no worse on that issue than any other measure, and gestures are particularly well suited to capturing visual-spatial representations.

Results

Figure 5 presents the proportion of gestures that were manipulative and configural broken down by time (first vs. last maneuver within each scenario) for each expertise group. We see the same pattern of results of change with time within each expertise group: a decrease in the proportion of configural gestures and an increase in the proportion of manipulative gestures. This

pattern is exactly what was predicted by the neurocomputational account: participants would go from a representation that is appropriate for times of high uncertainty about location (configural) to a representation that is appropriate for times of lower uncertainty about location (manipulative).

As there were no gestures about the 2D input in this situation, part of what the embodied problem solving perspective would predict did not come true. One could argue that the presence of configural representations, especially in early problem solving episodes, is consistent with the embodied problem solving focus on the external reality. It is interesting that the configural gestures relative to manipulative gestures were the lowest in the experts, suggesting an especially strong movement away from external reality in experts.

Of course, all of these conclusions are very tentative, as we have only examined performance in one situation and the results can be partially explained by each of the camps (not to mention various other ad hoc possible explanations of this simple pattern). It will be important to examine spatial representations in other tasks to see whether the neurocomputational perspective provides genuine insight.

Study 2: Expert/Novice Comparisons in Modern Submarining and fMRI Data Analysis

Overview

This study followed (group 1) cognitive neuroscientists at different expertise levels analyzing Functional Magnetic Resonance Imaging (fMRI) data and (group 2) submarine experts doing similar problem solving as in Study 1, but with a more modern interface that better affords display-based problem solving. The purpose of group 1 was to see whether we could predict representation choice in a very different domain, with a small rather than large external reality,

for example. The purpose of group 2 was to explore what role external input had on problem solving by using a different external input for the same basic task as in Study 1.

In the Submarine domain, we had problem solvers go through one complex scenario, as in Study 1. In the fMRI domain, we observed experts, intermediates, and novices analyzing their own data. In both domains, after 30-60 minutes of problem solving, we then stopped the data analysis activities, and showed the problem-solvers several one-minute videotape segments of their problem solving and asked them to explain what they knew and didn't know at that point in time, so that we could examine how they were representing their data spatially and what their uncertainty levels were. We examined the speech and gestures produced by problem solvers during those cued recall segments to measure their uncertainty levels and the way they represented their data spatially (acknowledging all along the potential dangers of relying on retrospective reports to measure internal representations). We then looked at uncertainty levels and representation choice as a function of each other as well as time and expertise.

fMRI Domain

The goal of fMRI is to discover both the location in the brain and the time course of processing underlying different cognitive processes. Imaging data is collected in research fMRI scanners hooked to computers that display experimental stimuli to their human subjects. Generally, fMRI uses a subtractive logic technique, in which the magnetic activity observed in the brain during one task is subtracted from the magnetic activity observed in the brain during another task, with the assumption that the resulting difference can be attributed to whatever cognitive processes occur in the one task but not the other. Moreover, neuronal activity levels are not directly measured, but rather one measures the changes in magnetic fields associated with oxygen-rich blood relative to oxygen-depleted blood. The main measured change is not the

depletion due to neuronal activity but rather the delayed over-response of new oxygen-rich blood moving to active brain areas, and the delay is on the order of 5 seconds, with the delay slightly variable by person and brain area. Data is analyzed visually by superimposing color-coded activity regions over a structural image of the brain (see Figure 6a), looking at graphs of mean activation level by region and/or over time (see Figure 6b) or across conditions (see Figure 6c), or looking at tables of mean activation levels by region across conditions (see Figure 6d). Elaborate, multi-stepped, semi-automated computational procedures are executed to produce these various visualizations, and given the size of the data (gigabytes per subject), many steps can take up to several minutes per subject. Inferential statistical procedures (e.g., t, ANOVA) are applied to confirm trends seen visually. Note that, as in the submarine domain, the input displays are very 2-dimensional, even though the underlying reality (activation in brain regions) is 3-dimensional. Unlike the submarine domain, however, the underlying reality takes place in a very small space (smaller than a breadbasket, relatively nearby) whereas in the submarine domain, the real space is many miles in every direction, with objects being the size of medium-sized buildings.

More Realistic Submarine Interface

While the basic task of finding other submarines using passive sonar remains fundamentally the same very difficult task, computational algorithms and visual displays designed to help the submariner have improved significantly. Figure 7 presents the more realistic interface that used in Study 2. It runs on a high-end Windows© personal computer, and is an unclassified simulation environment used in engineering development and training situations. It closely mirrors the actual displays used in modern US Navy submarines. Explaining all the displays found in Figure 7 is beyond the scope of this chapter, but suffice it to say that it includes both

egocentric and geosituational views, as well as alphanumeric best-guesses on target location, and that it includes explicit representations about the uncertainty in possible values of angle, distance, course, and speed of the target. Thus, in contrast to the Ned simulation used in Study 2, this environment affords better displayed-based problem solving, and thus we may see more display-based representations of space than in Study 1.

Participants

Submarine. There were 5 submarine experts who participated in Study, with similar expertise levels as the experts in Study 1.

fMRI. There were 10 fMRI participants, ranging from beginning graduate students to postdoctoral researchers. This study focused on naturalistic analysis of data, and faculty in this domain tend not to be directly involved in analysis of fMRI data, and instead work with students and postdocs after analyses have been carried out. We divided the participants into three expertise levels based on the number of studies they had carried out: 4 participants classified as Experts had carried out 4 or more fMRI studies, 4 participants classified as Intermediate has carried out between 2 and 3 studies, and 2 participants classified as Novices had carried out only 1 study. Since postdocs in this domain typically had earned their PhD with a technique other than fMRI, not all the postdocs were classified Experts and some of the graduate students were classified Experts. Although our fMRI Experts did not have the 10 years of focused practice that is typically required for world-class expertise, we are interested in expertise as a relative continuum, not as absolute categories.

Coding

The coding of gestures in Study 2 followed a similar procedure as in Study 1, although in this case we focused on gestures made during the various 'interesting minutes' cued responses rather

than on just first and last maneuvers, and we coded much more prevalent display-based gestures. Display-based gestures are gestures that described spatial relations in the discussed data, but occurred in a flat vertical (usually fronto-parallel) space, in contrast to manipulative gestures, which also took place in nearby space but gestured with 3-dimensional depth in object placement and/or size and shape, and in contrast to configural gestures, in which the hands were not representing the objects themselves but were merely pointers to objects off in a distance space. Figure 8b presents an example display-based gesture in which the participant takes about brain activation of two different spatial regions in terms of a flat bar-graph representation spatial region being represented one-dimensionally on the x-axis. By contrast, Figure 8a shows what a manipulative gestures looks like in this domain.

As in Study 1, we coded for uncertainty gestures (like shrugs and hand wiggles), but do not focus on those results here. Other gestures that were coded but not included in the current analyses were metaphorical gestures (in which space represented non-spatial dimensions like time), beating gestures (which simply keep time with speech or indicate points of emphasis in speech), and deictic gestures (point to the screen or a notebook on a desk, which is ambiguous about underlying spatial representations).

Predictions

As in Study 1, the symbolic perspective does not make obvious predictions—the adopted representation, especially by experts, could be anything, and all will depend upon what representations best support problem solving. The embodied problem-solving perspective makes the following predictions. First, fMRI scientists should use manipulative (real-world) and display-based (input) representations. Second, submariners should use configural (real-world) and display-based (input) representations. The neurocomputational perspective makes different

predictions. In fMRI, the end goal is not precise location, so the problem solvers should move to less precise representations (e.g., display-based representations). In submarining, the end goal is precise location, and thus the problem solvers should move to more precise representations (e.g., manipulative).

Results

Domain Differences in Expert Representations

Because we only collected data from experts in the submarine domain, to properly compare domain differences, we must focus on the expert data in the fMRI domain as well for a comparison across domains. Accordingly, Figure 9 presents the number of configural, display, and manipulative gestures for experts only in the fMRI and submarine domains.

Comparing the two domains, we can suggest several conclusions about expert representations. First, the underlying reality appears to matter a little. There were no configural gestures in the fMRI domain (to a large or distant brain) but there were some (although relatively few) configural gestures in the submarine domain. Second, the interface appears to matter. There were many display-based gestures in both domains, reflecting the input problem solvers received on the screen. Moreover, comparing to the results from Study 1, changing the interface to a more modern interface appears to impact the experts in that we now see a significant presence of display-based gestures. Third, the data from the submarine domain suggest that neurocomputational factors appear to matter a lot, because the most common representation (manipulative) corresponds to neither input nor external reality.

The diversity of representations within each group suggest that an account like ACT-R/S, in which there can be multiple spatial representations, is useful for highlighting representational variability. It is also the case that some participants used few spatial gestures overall. We do not

think they were not thinking spatially, but rather there are large individual differences in how much and what type of gestures people use. The majority who used at least three gestures had both manipulative and display gestures, suggesting the diversity does reside within individuals rather than reflecting individual choice of a single representation to use throughout problem solving.

Expertise Effects of Representation

Focusing in on the fMRI data, we can now turn to differences in preferred representation type as a function of expertise. Figure 10 presents the ratio of display to manipulative gestures (large numbers indicate relatively more display gestures). We can use this ratio in this domain because there were no configural gestures. We see a gradual increase in the use of display rather than manipulative representations with expertise. This difference is consistent across participants: 3/4 experts use more display than manipulative gestures, whereas 0/4 intermediate and 0/2 of novices use more displays than manipulative gestures).

Were these representation preferences held throughout problem solving, indicating that experts 'saw' different things in their data from the start, or was there a more complex pattern over time? We divided the cued minutes for each participant into early and late minutes. Unpacked by early/late, we see that experts start out the scenario with manipulative gestures but move to display-based gestures (see Figure 11). Thus, experts, like intermediates and novices, begin data analysis thinking about a three-dimensional brain (even though they are literally seeing 2-D slices of a 3-D brain). With problem solving, experts, unlike intermediates and novices, are better able to move to a more abstract 2D spatial representation: in the end, their question is not where in the 3D brain were there regions of activity, but rather how did functional

regions in the brain (which are more easily compressed into a single ordinal dimension) differ in their activity levels (by task or by time).

General Discussion

The goals of this chapter were to draw attention to a major weakness in theorizing in cognitive science (how can we predict representation choice), providing a new theoretical framing of the issue (by drawing out and contrasting predictions from the three major theoretical camps in cognitive science), and to provide some initial examinations of real world cognition in a complex domain to see how well the various predictions bear out.

Although the evidence is currently from only a two cases and a small number of participants, our data suggest the following directions. First, it appears that the external world (reality and input) does have some influence on internal representation choice. Moreover, it appears that reality primarily matters in novices and early in problem solving. Second, expert representations are best predicted by the match of task goals to neurocomputational constraints—experts appear to exapt particular, existing visual/spatial systems for problem solving on the basis of how well the computational abilities of those systems support the current needs/features of the given task. In particular, we have shown how spatial informational uncertainty is related to the selection of internal visual/spatial representations.

Of all the areas of psychology, research in complex, real-world problem solving seems most removed from all the excitement and breakthroughs in cognitive neuroscience of the last 15 to 20 years. This lack of change in research on higher-level cognition is not arbitrary or representative of stubbornness by a particular research community. Instead, it reflects the difficulties in bring neuroscience methodologies to studying something so complex as higher-level cognition, which almost by definition, influences the integration of many brain regions and brain systems in

complex ways. We hope that the work described in this chapter can show a different way in which neuroscience can bring new insights to the study of higher-level cognition: bringing in theoretical constraints on core components of the problem-solving system based on neuroscience data and theories. We hope that we have also made some progress in convincing researchers of complex cognition that we need to move beyond relying solely on our old theoretical friends of task structure, memory constraints, and embodied cognition to understand complex problem solving.

Caveats

It is important to acknowledge that the current story is just the beginning of the story. Much further empirical work must be done to establish the value of the current story over various alternative explanations of our presented data. As we argued in the beginning of the chapter, the measurement problem for internal representations is a very difficult one. Consequently, we do not know for sure that gestures cleanly correspond to internal representations. Instead, the representations that we observed might only correspond to a subset of the representations that the problem solvers were entertaining, and perhaps the subset that was easiest to communicate to the listener. Moreover, the act of communication may drive representation choice more than the basic task itself, and the pragmatics of spatial communication by gesture may be important here. Further work with other measures of internal representations, in addition to collecting more data from more participants and in more domains, should be done to strongly validate the story that we are telling about our data.

Contributions of Different Perspectives—Building the Computational Story

What have the different perspectives contributed to our current understanding of how problem solvers choose internal representations? We argue that each perspective has built upon

the previous, elaborating the computational story of cognition. The symbolic perspective began by showing us that computational rather than physical properties per se matter—the structure of the problem space matters much more than the particular physical device with which we interact. The embodied cognitive perspective has shown us that many of our computations are performed on external objects or are grounded in knowledge about the world, so input and reality matters in specifying the nature of the computations. Finally, the neurocomputational perspective has shown us that our choice of representations and their computational properties are strongly influenced by our neurobiology. Thus, a complete computational account of problem solving in a domain includes the task, the environment, and the computational abilities of the problem solver.

Back to Uncertainty in Data Analysis

Linking back to data analysis and uncertainty themes in this volume, our work suggests that uncertainty has perhaps more roles in problem solving that others have discussed. First, it is an object in itself to detect. Uncertainty varies across situations, and the problem solver needs to be able to detect the situations in which uncertainty levels are especially high. It is this role of uncertainty that much work in statistics lies (including the work in this volume). Second, uncertainty is an object to problem solve about. When one moves into real problem solving applications, uncertainty has valence (i.e., it is bad for problem solving), and the problem solver must engage in activities to reduce uncertainty. The work by Trickett et al. in this volume discusses this aspect of uncertainty. Third, uncertainty is an object that influences basic representation choice, and that basic representation choice will influence many other aspects of problem solving. It is this third role that has perhaps not been discussed previously, although we suspect it maybe an interesting lens even in basic statistics course problem solving.

Acknowledgements

Work on this project was supported by grants Grant N00014-02-1-0113 and N00014-03-1-0061 to the first author from the Office of Naval Research.

References

Anderson, J. R. (1990). *The adaptive character of thought*. Hillsdale, NJ: Erlbaum.
Anderson, J. R., Bothell, D., Byrne, M., & LeBiere, C. (2002). An integrated theory of the mind. from http://act-r.psy.cmu.edu/papers/403/IntegratedTheory.pdf
Anderson, J. R., & Lebiere, C. (1998). *Atomic components of thought*. Mahwah, NJ: Erlbaum.
Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral & Brain Sciences, 22*(4), 577-660.
Broadbent, D. E. (1957). A mechanical model for human attention and immediate memory. *Psychological Review, 64*, 205-215.
Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5*, 121-152.
Dietrich, E., & Markman, A. B. (2000). Cognitive dynamics: Computation and representation regained. In E. Dietrich & A. B. Markman (Eds.), *Cognitive dynamics* (pp. 5-30). Mahwah, NJ: Erlbaum.
Ehret, B. D., Gray, W. D., & Kirschenbaum, S. S. (2000). Contending with complexity: Developing and using a scaled world in applied cognitive research. *Human Factors, 42*(1), 8-23.
Fu, W.-T., & Gray, W. D. (2004). Resolving the paradox of the active user: Stable suboptimal performance in interactive tasks. *Cognitive Science, 28*(6).
Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.
Gobet, F., & Simon, H. A. (1996). Recall of random and distorted chess positions: Implications for the theory of expertise. *Memory & Cognition, 24*(4), 493-503.
Gould, S. J., & Vrba, E. S. (1982). Exaptation--a missing term in the science of form. *Paleobiology, 8*, 4-15.
Gray, W. D., John, B. E., & Atwood, M. E. (1993). Project ernestine: Validating goms for predicting and explaining real-world task performance. *Human Computer Interaction, 8*(3), 237-309.
Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology, 160*, 106-154.
Hutchins, E. (1995a). *Cognition in the wild*. Cambridge, MA, USA: MIT Press.
Hutchins, E. (1995b). How a cockpit remembers its speeds. *Cognitive Science, 19*(3), 265-288.
Just, M. A., & Carpenter, P. A. (1976a). Eye fixations and cognitive processes. *Cognitive Psychology, 8*(4), 441-480.
Just, M. A., & Carpenter, P. A. (1976b). The role of eye-fixation research in cognitive psychology. *Behavior Research Methods, Instruments, & Computers, 8*(2), 139-143.
Kaplan, C. A., & Simon, H. A. (1990). In search of insight. *Cognitive Psychology, 22*, 374-419.
Kieras, D. E., & Meyer, D. E. (1997). An overview of the epic architecture for cognition and performance with application to human-computer interaction. *Human-Computer Interaction, 12*(4), 391-438.
Klahr, D., & Dunbar, K. (1988). Dual space search during scientific reasoning. *Cognitive Science, 12*, 1-48.
Kosslyn, S. M. (1994). *Image and brain: The resolution of the imagery debate*. Cambridge. MA: MIT Press.

Kosslyn, S. M., Ball, T. M., & Reiser, B. J. (1978). Visual images preserve metric spatial information: Evidence from studies of image scanning. *Journal of Experimental Psychology: Human Perception and Performance, 4*, 46-60.

Kosslyn, S. M., Pascual-Leone, A., Felican, O., Camposano, S., Keenan, J. P., Thompson, W. L., et al. (1999). The role of area 17 in visual imagery: Convergent evidence from pet and rtms. *Science, 284*(April 2), 167-170.

Kosslyn, S. M., Pinker, S., Smith, G., & Shwartz, S. P. (1979). On the demystification of mental imagery. *Behavioral and Brain Science, 2*, 535-548.

Kosslyn, S. M., Thompson, W. L., Kim, I. J., & Alpert, N. M. (1995). Topographical representations of mental images in primary visual conrtex. *Nature, 378*(Nov 30), 496-498.

Kotovsky, K., Hayes, J. R., & Simon, H. A. (1985). Why are some problems hard? Evidence from tower of hanoi. *Cognitive Psychology, 17*(2), 248-294.

Larkin, J. H., McDermott, J., Simon, D., & Simon, H. (1980). Expert and novice performance in solving physics problems. *Science, 208*, 140-156.

Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth 10,000 words. *Cognitive Science, 4*, 317-345.

Lovett, M. C., & Schunn, C. D. (1999). Task representations, strategy variability and base-rate neglect. *Journal of Experimental Psychology: General, 128*(2), 107-130.

Markman, A. B. (1999). *Knowledge representation*. Mahwah, NJ: Erlbaum.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago, IL, USA: University of Chicago Press.

Neisser, U. (1976). *Cognition and reality: Principles and implications of cognitive psychology*. San Francisco: W. H. Freeman.

Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.

Perkins, D. N. (1994). Insight in minds and genes. In R. S. Sternberg & J. E. Davidson (Eds.), *The nature of insight*. Cambridge, MA: MIT Press.

Previc, F. H. (1998). The neuropsychology of 3-d space. *Psychological Bulletin, 124*(2), 123-164.

Pylyshyn, Z. W. (1973). What the mind's eye tells the mind's brain: A critique of mental imagery. *Psychological Bulletin, 80*, 1-24.

Pylyshyn, Z. W. (1981). The imagery debate: Analogue media versus tacit knowledge. *Psychological Review, 88*, 16-45.

Pylyshyn, Z. W. (1994). Some primitive mechanisms of spatial attention. *Cognition, 50*(1-3), 363-384.

Pylyshyn, Z. W. (2002). Mental imagery: In search of a theory. *Behavioral and Brain Science, 25*(2), 157-182.

Scaife, M., & Rogers, Y. (1996). External cognition: How do graphical representations work? *International Journal of Human-Computer Studies, 45*(2), 185-213.

Schunn, C. D., & Klahr, D. (1996). *The problem of problem spaces: When and how to go beyond a 2-space model of scientific discovery*. Paper presented at the The 18th Annual Conference of the Cognitive Science Society, San Diego, CA.

Schunn, C. D., & Klahr, D. (2000). Discovery processes in a more complex task. In D. Klahr (Ed.), *Exploring science: The cognition and development of discovery processes*. Cambridge, MA: MIT Press.

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review, 63*, 129-138.

Suchman, L. A. (1987). *Plans and situated action: The problem of human-machine communication*. New York: Cambridge University Press.

Tabachneck-Schijf, H. J. M., Leonardo, A. M., & Simon, H. A. (1997). Camera: A computational model of multiple representations. *Cognitive Science, 21*(3), 305-350.

Treisman, A. M. (1969). Strategies and models of selective attention. *Psychological Review, 76*(3), 282-299.

Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. In D. J. Ingle, M. A. Goodale & R. J. W. Mansfield (Eds.), *Analysis of visual behavior*. Cambridge, MA: MIT Press.

Wolfe, J. M. (1994). Guided search 2.0: A revised model of visual search. *Psychonomic Bulletin and Review, 1*(2), 202-238.

Table 1. Comparison of general predictions about representational choice from each the three theoretical camps.

Theoretical Camp	Use Affordances?	External Matters?	Internal Choices?
Symbolic	√	Maybe	Anything
Embodied	√	Yes	Anything
Neurocomputational	√	Aspects that are processed	Fixed set, Exaptation

Figure 1. Three visual/spatial representation systems posited in ACT-R/S, the size and location of space they cover, and the basic tasks they typically support.

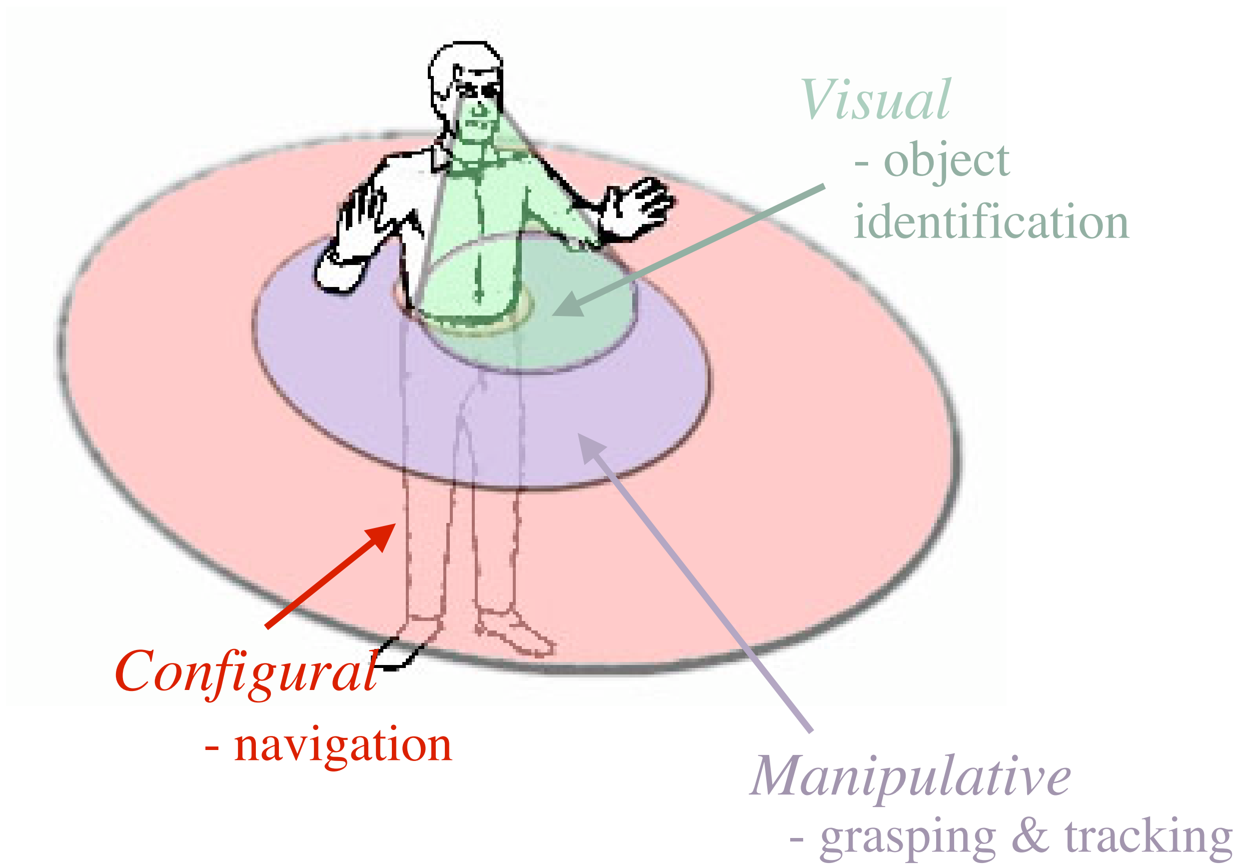
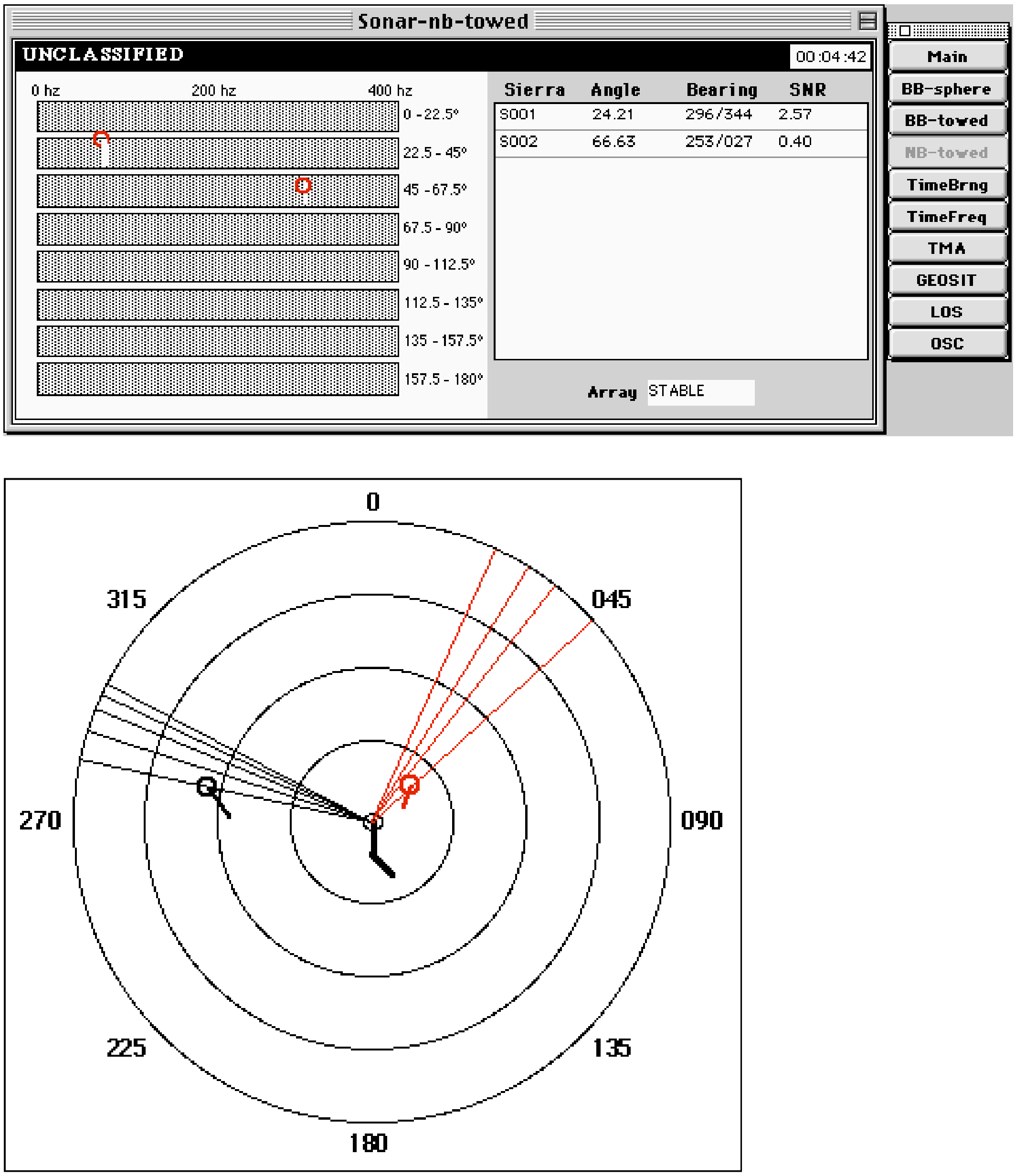


Figure 2. Two sample screen shots from the Ned submarine simulation environment used in Study 1.



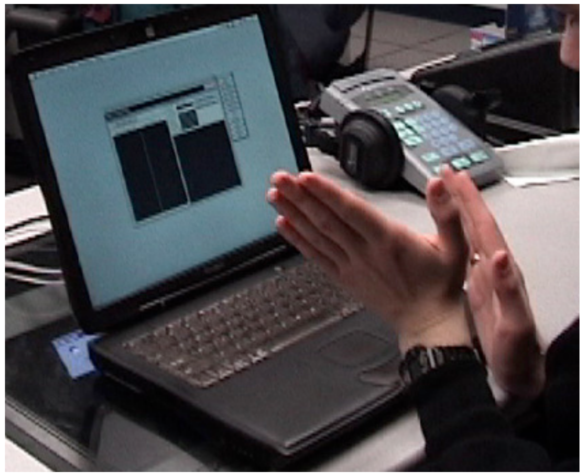


Figure 3**.** A participant's configural gesture produced during his think-aloud protocol "…bearing around course oh three five, our own-ship course is about three five seven, we'll be about…here".

Figure 4. A participant's manipulative gesture produced during a hotwash, saying *“I should’ve gone left…come left and gone be**hind** him…”*.

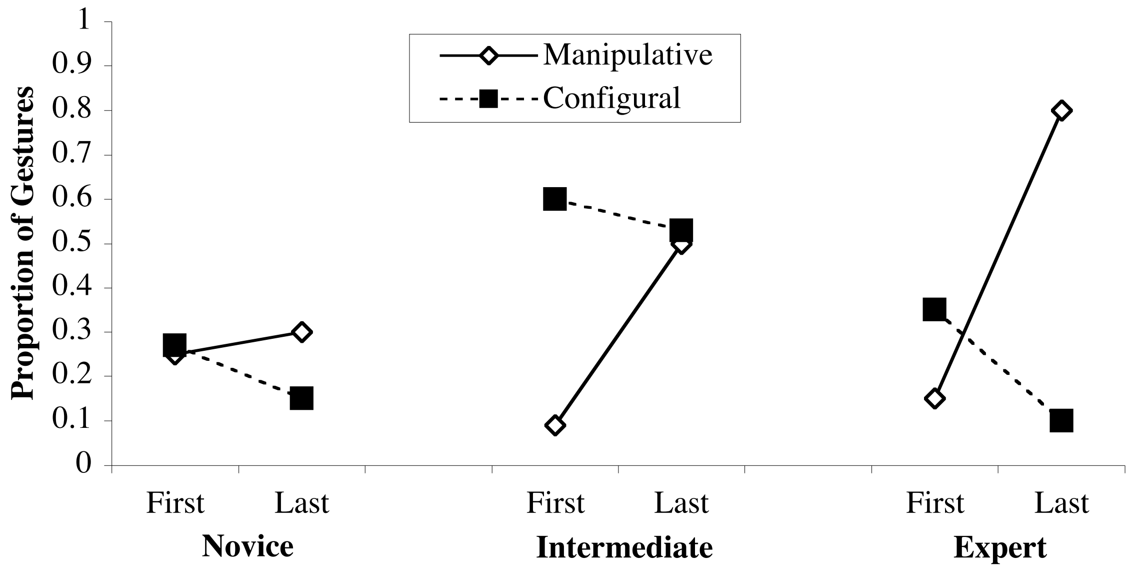
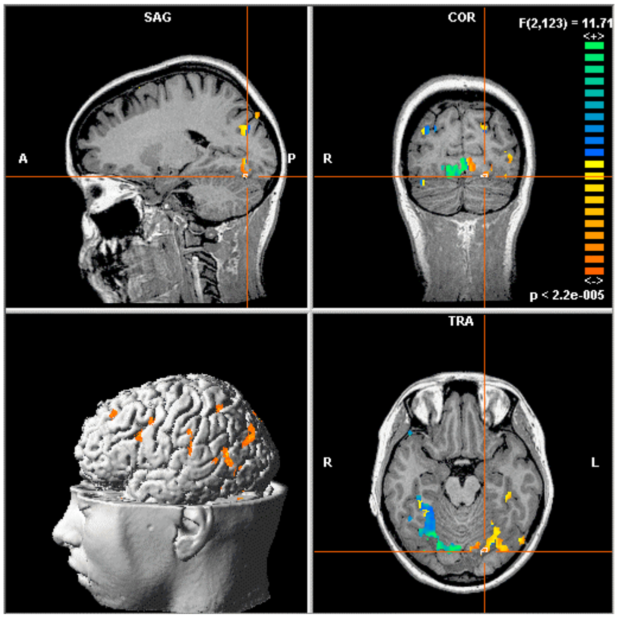
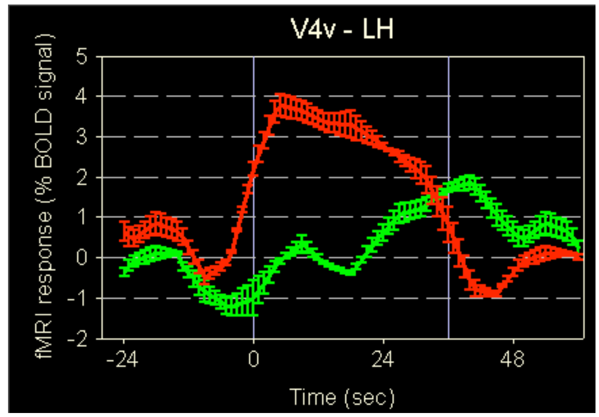
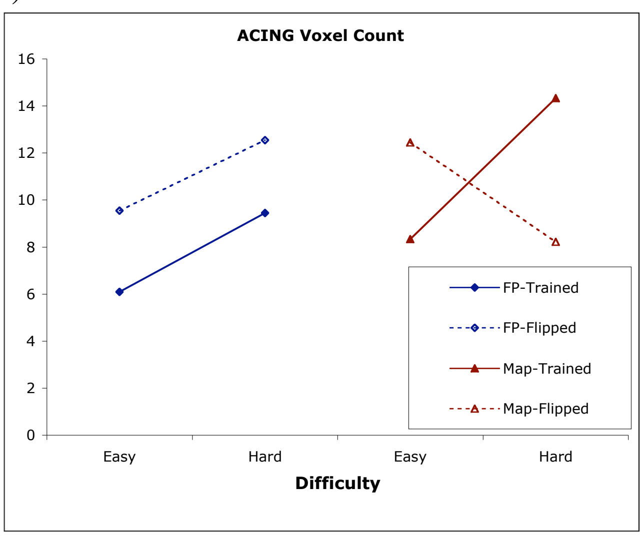


Figure 5. Proportion of gestures that were manipulative and configural gestures for the first and last maneuver of each scenario for novice (students), intermediates (instructors), and experts (commanders) in Study 1.





c)



d)

	TrainedEasy	TrainedHar	FlippedEasy	FlippedHard
ROI	Count	Count	Count	Count
1	2	3	4	5
ACING	8.33	14.33	12.44	8.22
CALC	72.33	87.56	75.89	103.22
LCBELL	21.11	33	27.22	28.89
RCBELL	19.44	38.44	28.33	33.67
LCBELL+RCBELL	40.56	71.44	55.56	62.56
LDLPFC	**32.78**	**51**	**38.22**	**40.56**
RDLPFC	**37.56**	**61.67**	**43.44**	**57.33**
DLPFC	**70.33**	**112.67**	**81.67**	**97.89**
LFEF	5.33	8.22	4.67	7.33
RFEF	3.44	5.56	3.78	6.56
LFEF+RFEF	8.78	13.78	8.44	13.89
LHG	12.67	1.44	5.56	1.56
RHG	14	3	6	3.78
LHG+RHG	26.67	4.44	11.56	5.33
LIES	77.56	104.89	84.78	109.44
RIES	59.11	76.56	59.56	81.67
LIES+RIES	136.67	181.44	144.33	191.11
LIPL	26.44	22.22	27.33	20.67
RIPL	19.11	21.89	20.33	20.56
LIPL+RIPL	45.56	44.11	47.67	41.22
LIPS	**54.67**	**76.56**	**61**	**75.22**
RIPS	**37.67**	**64.11**	**48.11**	**72.56**
IPS	**92.33**	**140.67**	**109.11**	**147.78**

Figure 6. Kinds of visualizations examined in analysis of fMRI data: a) degree of activation indicated with a color scale superimposed over a gray-scale structural brain image in three different planar slices and a surface cortex map; b) graph of percent signal change in a brain region as a function of time relative to a stimulus presentation in two different conditions (red and green); c) graph of number of activated voxels in an area as a function of various condition manipulations; and d) table of number of activated voxels in different brain areas (Regions of Interest) as a function of different conditions.

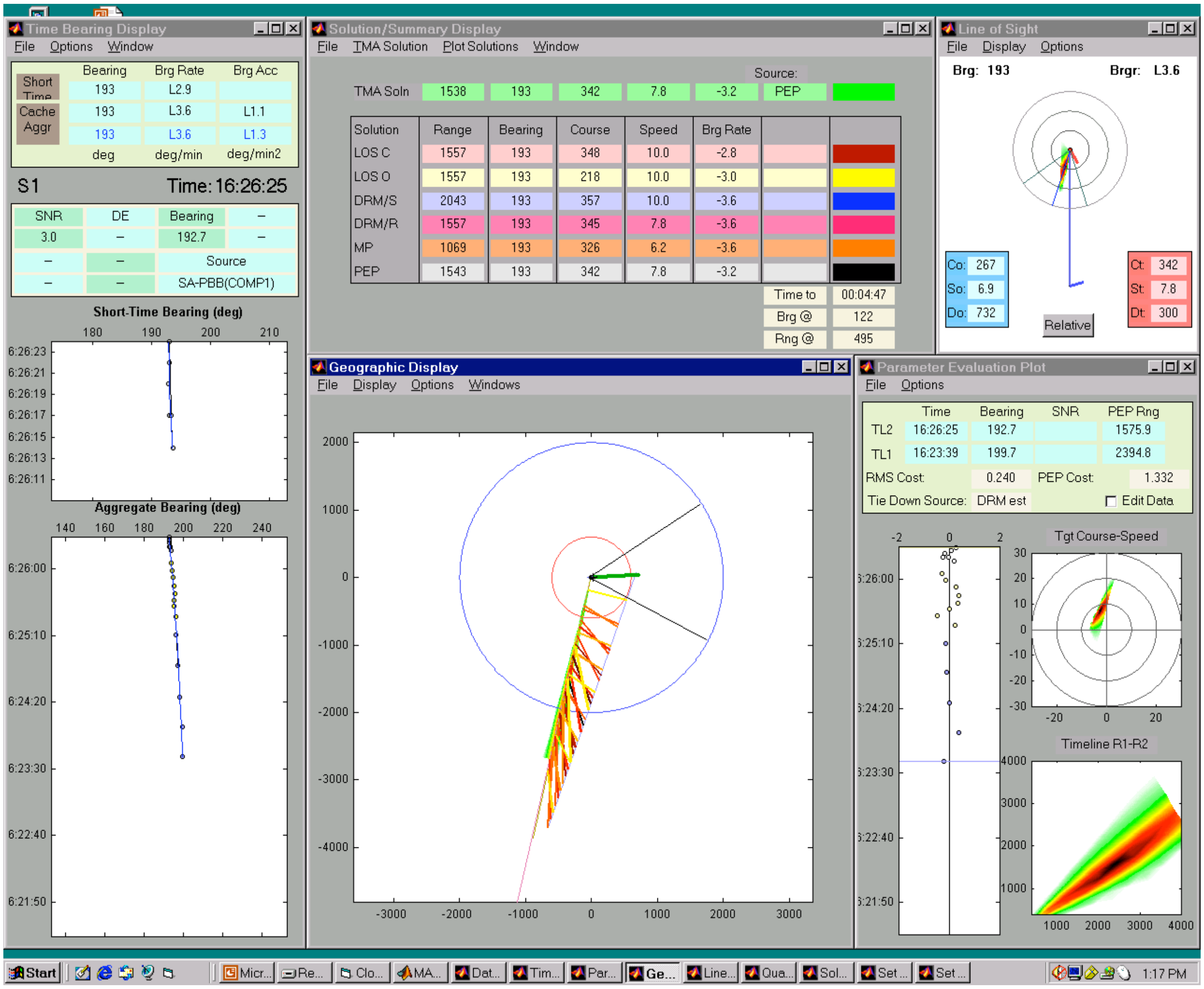


Figure 7. Modern submarine display used in Study 2.

a) b)

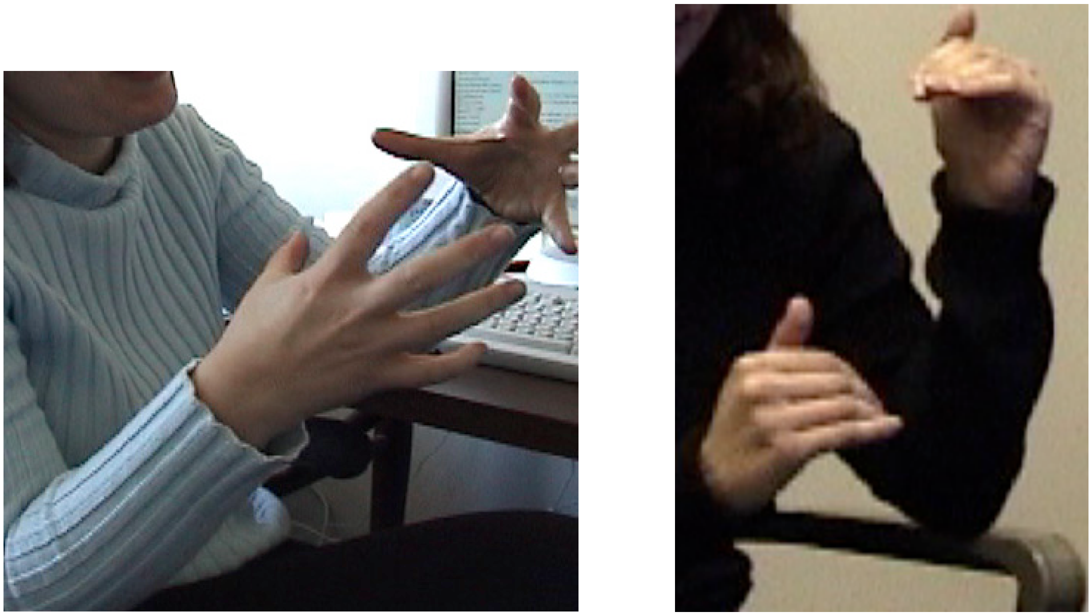


Figure 8. Example spatial gestures from the fMRI domain. a) a manipulative gesture, "… if you have, like, this massive thing, the peak is really in there…", and b) an example display-based gesture, "...I found out that, it looked like there's a difference between frontal and hippocampal activation..."

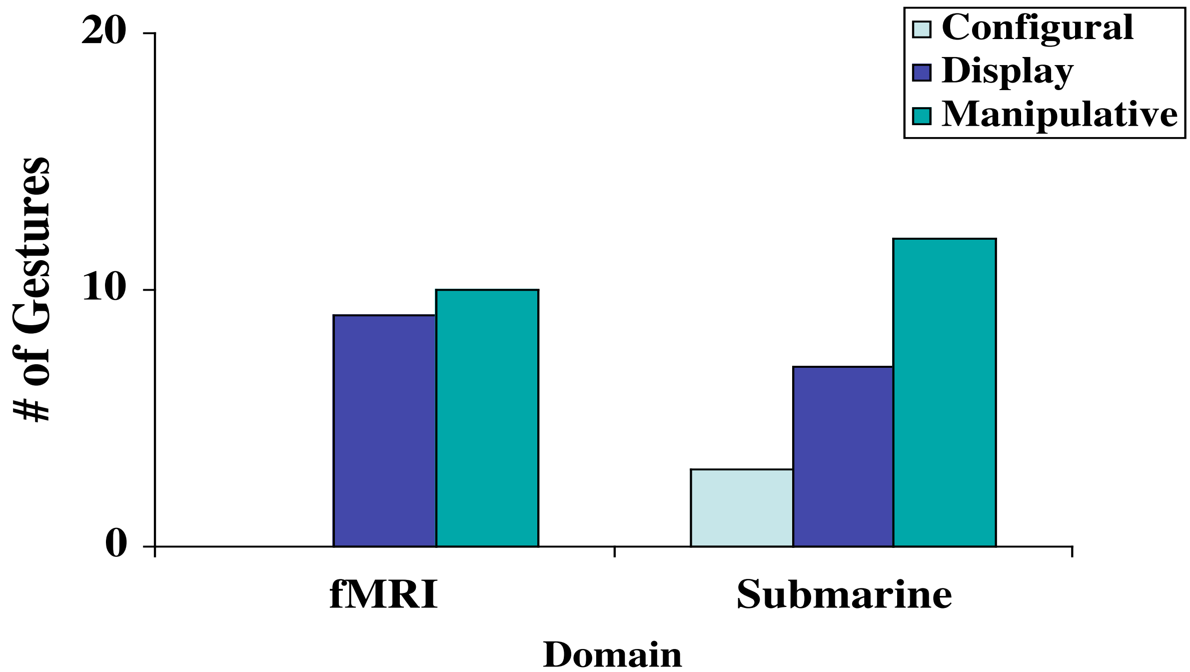


Figure 9. For experts only in study 2, the number of configural, display, and manipulative gestures found in each domain.

Figure 10. For fMRI scientists in Study 2, the ratio of display to manipulative gestures in each expertise group.

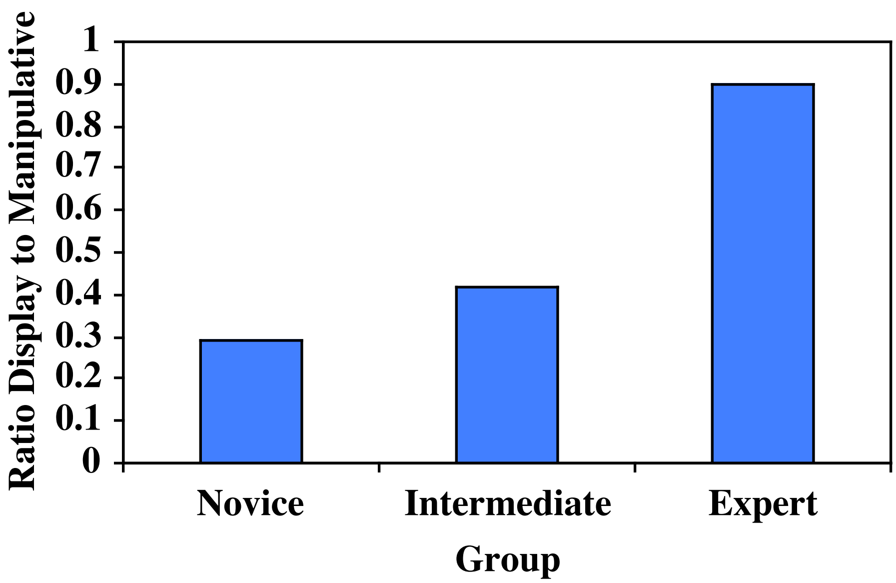


Figure 11. For fMRI scientists in Study 2, the ratio of display to manipulative gestures in each expertise group, split by the first half vs. second half of cued minutes.